

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>5</sup> :</b> C12P 21/02, C12N 15/00 C07H 21/04	<b>A1</b>	<b>(11) International Publication Number:</b> WO 90/12107 <b>(43) International Publication Date:</b> 18 October 1990 (18.10.90)
<b>(21) International Application Number:</b> PCT/US90/01738 <b>(22) International Filing Date:</b> 2 April 1990 (02.04.90) <b>(30) Priority data:</b> 332,632 31 March 1989 (31.03.89) US <b>(60) Parent Application or Grant</b> <b>(63) Related by Continuation</b> US 332,632 (CIP) Filed on 31 March 1989 (31.03.89) <b>(71) Applicant (for all designated States except US):</b> THE SALK INSTITUTE BIOTECHNOLOGY/INDUSTRIAL AS- SOCIATES, INC. [US/US]; 10280 North Torrey Pines Road, La Jolla, CA 92037 (US).	<b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only) :</b> FITZMAURICE, Le- ona, Claire [US/US]; 10369 Caminito Surabaya, San Diego, CA 92131 (US). MIRKOV, Theodore, Erik [US/ US]; 12202 Creswick Court, San Diego, CA 92138 (US). <b>(74) Agents:</b> WATT, Phillip, H. et al.; Fitch, Even, Tabin & Flannery, Room 900, 135 South LaSalle Street, Chicago, IL 60603 (US). <b>(81) Designated States:</b> AT (European patent), BE (European + patent), CA, CH (European patent), DE (European pa- tent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European pa- tent), SE (European patent), US.  <b>Published</b> <i>With international search report.</i>	
<b>(54) Title:</b> RECOMBINANT EXPRESSION SYSTEM BASED ON SATELLITE TOBACCO MOSAIC VIRUS  <b>(57) Abstract</b>  The present invention provides the sequence of the RNA genome of satellite tobacco mosaic virus. Based on said sequence and the discovery that heterologous RNA can be accommodated in said genome without eliminating replicability, the invention provides compositions comprising modified genomes of the virus, or cDNAs of such genomes, to transform plant cells in vitro and in vivo to make desired RNAs or proteins. Further, the invention provides methods and intermediates for making such compositions and methods of using such compositions.		

### DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

#### *FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MR	Mauritania
BE	Belgium	GA	Gabon	MW	Malawi
BF	Burkina Faso	GB	United Kingdom	NL	Netherlands
BG	Bulgaria	HU	Hungary	NO	Norway
BJ	Benin	IT	Italy	RO	Romania
BR	Brazil	JP	Japan	SD	Sudan
CA	Canada	KP	Democratic People's Republic of Korea	SE	Sweden
CF	Central African Republic	KR	Republic of Korea	SN	Senegal
CG	Congo	LI	Liechtenstein	SU	Soviet Union
CH	Switzerland	LK	Sri Lanka	TD	Chad
CM	Cameroon	LU	Luxembourg	TG	Togo
DE	Germany, Federal Republic of	MC	Monaco	US	United States of America
DK	Denmark				

/  
RECOMBINANT EXPRESSION SYSTEM BASED ON  
SATELLITE TOBACCO MOSAIC VIRUS

## TECHNICAL FIELD

5           The present invention relates to genetic engineering of plants and, more particularly, to such engineering utilizing the properties of RNA plant viruses.

10           Thus, the invention concerns satellite tobacco mosaic virus (STMV), recombinant STMV RNA molecules containing exogenous RNA segments, which are heterologous to naturally occurring STMV RNAs, and a recombinant expression system making possible production of a desired gene product in the cytoplasm of a plant  
15 infected with recombinant STMV RNA molecules and a helper virus.

          In one aspect, the present invention provides infectious recombinant RNA molecules derived from satellite tobacco mosaic virus (STMV) ssRNA.

20           In another aspect, the present invention concerns DNA transcription vectors containing substantially full-length cDNA copies of infectious STMV ssRNA.

25           In a further aspect, the invention relates to infectious recombinant RNA molecules derived from STMV ssRNA, having an exogenous RNA segment at a site that is non-essential for RNA replication in a host cell.

30           In a still further aspect, the present invention relates to DNA transcription vectors containing substantially full-length cDNA copies of such infectious recombinant RNA molecules.

35           The present invention further concerns a method of transforming plant cells by introducing into the cytoplasm of such cells infectious recombinant RNA molecules derived from STMV ssRNA, having such exogenous RNA segments, and a helper virus of STMV.

          The present invention also provides a method for the production of exogenous proteins in the cytoplasm

of plant cells by utilizing such genetically manipulated infectious RNA molecules, or corresponding cDNAs, and a helper virus of STMV.

5 BACKGROUND OF THE INVENTION

Members of several different plant virus groups support the replication of another virus (Kassanis, Intervirology 15, 57-70 (1981); Buzen et al., Phytopathology 74, 313-8 (1984); Gingery and Louie, 10 Phytopathology 75, 870-4 (1985)). Viruses which have a dependence on another virus for their replication were first reported by Kassanis, J. Gen. Microbiol. 27, 477-488 (1962) and are known as satellite viruses. Satellite viruses of plant viruses depend completely on a specific 15 helper virus (or viruses) for their replication but appear to share little, if any, nucleotide sequence similarity with their helpers (Francki, Ann. Rev. Microbiol. 39, 151-74 (1985); Murant and Mayo, Ann. Rev. Phytopathol. 20, 49-70 (1982)). Satellites may interfere 20 with helper virus synthesis or may modify disease symptom expression in the host plant. Satellite viruses differ from so-called satellite RNAs. A satellite virus encodes a capsid protein that specifically encapsidates its own satellite virus RNA. In contrast, a satellite RNA is 25 encapsidated by capsid protein provided by an helper virus RNA.

Satellite viruses include those of tobacco necrosis virus ((STNV); Kassanis, Intervirology 15, 57-70 (1981)), panicum mosaic virus ((SPMV); Buzen et al., 30 Phytopathology 74, 313-8 (1984); Masuta et al., Virology 159, 329-38 (1987)) and maize white line mosaic virus (Gingery and Louie, Phytopathology 75, 870-4 (1985)). The helper viruses of these satellite viruses are unrelated, but all of the helpers have monopartite 35 single-stranded RNA (ssRNA) genomes and exist as icosahedral 30nm particles. The satellite viruses all exist as 17nm icosahedral particles which are

antigenically unrelated to their helpers and to each other. Although STNV and SPMV have been well characterized and the nucleotide sequences of their genomes have been determined, there are no data on the genome sequences and genome organizations of their respective helpers.

Satellite tobacco mosaic virus (STMV) is the most recently identified plant satellite virus (Valverde and Dodds, J. Gen. Virology 67, 1875 (1986) and J. Gen. Virology 68, 965 (1987)). It consists of a single,  $0.3 \times 10^6$  Mr linear, messenger-sense, single-stranded RNA (ssRNA) and multiple copies of a 17,500 Mr capsid protein and depends on helper viruses for its replication. All of STMV's helper viruses are rod-shaped tobamoviruses, including tobacco mosaic virus (TMV), that is a prototype member of the group.

The structure of STMV, including the complete nucleotide sequence of its ssRNA genome, has been thoroughly investigated (E.Mirkov: cDNA Cloning, Nucleotide Sequences, in vitro Translation, and Genome Organization of STMV, Dissertation, University of California, Riverside (March 1988)). However, as will be described more specifically hereinafter, the published sequence of the STMV RNA contains numerous significant errors, particularly in the 3'- and 5'-termini, and the total length of the ssRNA genome as reported (1,065 ribonucleotides) is incorrect. The STMV ssRNA has two open-reading frames (ORFs). The first ORF encodes a 6,800 Mr protein that corresponds in size to a major in vitro translation product directed by STMV RNA. The second ORF encodes a 17,500 Mr protein that corresponds in size to the other major in vitro translation product synthesized under the direction of STMV RNA. The first 12 codons of this second ORF were found to correspond to the sequence of 12 N-terminal amino acids of the capsid protein. Western-blot analysis of these in vitro translation products revealed that the 17,500 Mr protein

is antigenically related to the authentic capsid protein, while the 6,800 Mr protein is not. Time course analysis of in vitro translation products demonstrated that the 6,800 Mr protein is synthesized at the same time as the capsid protein and does not, therefore, arise by the proteolytic cleavage of a larger precursor polypeptide, suggesting that the genome of STMV acts as a polycistronic, eukaryotic, messenger RNA.

The tobamoviruses, to which the known helper viruses of STMV belong, are known in, and available to, the art. Tobacco mosaic virus (TMV), for example, is one of the best characterized plant viruses from molecular biological and physicochemical points of view. For a review of this group see Molecular Plant Virology, Volume II, J. W. Davies, editor, CRC Press Inc., Boca Raton, Florida U.S.A. (1985). Known representatives of this group include the TMV strains U1, U2 and U5. These strains are described in Goelet et al., Proc. Natl. Acad. Sci. USA 78, 5818 (1982) and Garcia-Arenal, Virology 166, 201 (1988) and infectious clones of the U1 strain are available from Prof. William Dawson, University of California, Riverside.

The preparation of infectious RNA transcripts from cloned cDNAs of cucumber mosaic viral satellites is described in Whitmer et al., Biochemical and Biophysical Research Communications 135, 290 (1986). Complete cDNA copies of two variants of the satellite of cucumber mosaic virus, CARNA 5, were cloned in a transcription vector, and coinfectd with cucumber mosaic viral RNAs on tomato plants.

Meshi et al., Proc. Natl. Acad. Sci. USA 83, 5043 (1986) cloned full-length double-stranded cDNAs of tobacco mosaic virus (TMV) RNA into a transcription vector, and the in vitro RNA transcripts were used to infect tobacco plants.

Emmelo et al., Virology 157, 480 (1987) reported that mechanical inoculation of cowpea leaves

with cloned full-size copies of satellite tobacco necrosis virus (STNV) in the presence of its helper, tobacco necrosis virus (TNV), resulted in the appearance of infectious virus particles.

5           In vitro synthesis of infectious RNA copies of the virulent satellite (RNA C) of turnip crinkle virus (TCV) is disclosed in Simon et al., Virology 156, 146 (1987). Full-length cDNA copies of the satellite were inserted into an expression vector, and the RNA  
10 transcripts synthesized in vitro from the cDNAs were coinoculated with helper virus RNA into 3-week-old turnips. Plus-strand RNA copies of the satellite were found to be infectious, while minus-strand copies were not.

15           European Patent Application Publication No. 0 248 077 discloses a process for increasing protein production from a cDNA encoding an eukaryotic protein by joining to the 5'-terminus of the cDNA a nucleotide sequence which has a regulatory role, in expression of a  
20 coat protein of a plant virus, by increasing competitive activity and RNA translation leading to the coat protein. Suitable viruses as sources of the nucleotide sequence, which regulates expression of coat protein, are said in the published application to include alfalfa mosaic  
25 virus, brome mosaic virus, black beetle virus, turnip yellow mosaic virus, and satellite tobacco necrosis virus.

          European Patent Application Publication No. 0 194 809 discloses the modification of the genome of  
30 brome mosaic virus (BMV) to include an exogenous (i.e., heterologous) RNA segment, and the introduction of the modified RNA into a host cell, wherein the virus replicates and an exogenous gene product (i.e., a gene product heterologous to BMV) is produced. BMV is a  
35 multipartite RNA virus, i.e., the total genetic information required for replication and productive infection is divided into more than one discrete RNA

molecule. According to EPA 0 194 809, the cDNA corresponding to one of the RNA components (RNA3) of the BMV genome was modified by replacing a certain area (the RNA4 coding sequence) with a CAT gene (i.e., a cDNA encoding bacterial chloramphenicol acetyltransferase (CAT)). From the cDNAs corresponding to the RNA components of the virus genome, including the modified cDNA, transcripts were made, and barley seedlings were infected with a mix of the cDNA transcription products. The fact that the RNA4 coding sequence of the RNA3 component of BMV is not required for infectivity was disclosed by Lane et al., Nature 232, 40 (1971). In mixing experiments, the authors found that infectivity resulted when the combination of RNAs 1, 2, and 3 was used, and "the omission of any except component 4 from a mixture markedly decreases specific infectivity".

European Patent Application Publication No. 0,229,174 describes production of transgenic plants, that are able to express foreign proteins, by infection with a retroviral vector.

An expression system for making desired proteins in plants and based on satellite tobacco mosaic virus (STMV) has heretofore not been available or known. The structure and properties of STMV, including those which the present inventors have discovered as described hereinbelow, make the virus particularly apt as a target for genomic modification and associated use to prepare vectors for such an expression system. Until the present invention, such use of the virus and the advantages of such use could not be realized.

#### SUMMARY OF THE INVENTION

The present invention provides, for the first time, the correct complete nucleotide sequence of the genomic RNA of satellite tobacco mosaic virus (STMV). The sequence as earlier published (see Mirkov's Dissertation, supra) contained several incorrect



nucleotides at and near the 3'- and 5'-termini, and erroneously included extra nucleotides that resulted in a longer RNA molecule (1,065 nucleotides) than the viral RNA (1059 nucleotides) actually has. It has been found that the 3'-terminal nucleotides of STMV RNA and either TMV U1 or TMV U2/U5 show a great degree of sequence similarity, with two nearly identical regions of 40 and 50 bases, respectively. This is surprising in view of the earlier assumption that satellite viruses and their helpers share little, if any, genomic sequence similarity.

The invention further provides infectious recombinant RNA molecules that are derived from STMV ssRNA, preferably via transcription from a substantially full-length cDNA copy of the genomic STMV ssRNA or a modification thereof which incorporates a segment coding for an exogenous protein (i.e., a protein heterologous to those made by the naturally occurring STMV genome) desired to be expressed in a plant cell infected with the ssRNA molecules.

The present invention further relates to a recombinant expression system based on the use of STMV. The expression system allows for expression of a desired exogenous gene (i.e., one heterologous to the naturally occurring STMV genome), in the cytoplasm of a plant coinfecting with recombinant STMV RNA molecules of the invention and a helper virus. Because the infection remains cytoplasmic, the recombinant trait is not passed on to progeny plants and is thus easily contained in the plant. The expression system can, for example, be used to transfer economically attractive traits, e.g. disease resistance, to plants.

The expression system of the present invention provides the following major advantages over the already known systems:

1. The genome of STMV encodes a demonstrable in vivo protein product, a coat protein, which is not the

cas for some other currently popular model systems, such as the systems based on the satellite cucumber mosaic virus and tobacco ringspot virus. This allows for an easy assay of gene expression.

5           2. In contrast to satellite tobacco necrosis virus (STNV), which also encodes its coat protein, STMV is systemically distributed in the host in which it accumulates, as is its helper. This means that expression of a foreign gene in the host may well be  
10 throughout an entire plant.

3. The plant hosts used when working with this satellite are herbaceous, including many solanaceous plants, for which transformation systems are available.

15           4. The STMV virus is readily introduced into plants as virion particles or RNA by mechanical inoculation, and activation is with a helper virus, which is also mechanically transmitted.

20           5. The spherical particles of the satellite are readily distinguished from the rod-shaped particles of its helper.

25           6. The genome structure is unique when compared to other satellite viruses. These unique features include non-phosphorylated 5'-termini, the degree of similarity to its helper viruses, the genome organization, and the ability to function as a polycistronic mRNA.

30           In one aspect, the present invention relates to an infectious recombinant RNA molecule derived from satellite tobacco mosaic virus (STMV) ssRNA. Preferably, said RNA molecule is unaccompanied by endogenous components not essential for its infectious properties, and is preferably derived from an authentic STMV ssRNA molecule, having a nucleotide sequence as shown in Figure 1.

35           In a further aspect, the invention concerns a DNA transcription vector containing a substantially full-length cDNA copy of an infectious STMV ssRNA.

In a still further aspect, the present invention relates to infectious recombinant RNA molecules derived from STMV ssRNA, having an exogenous RNA segment at a site that is non-essential for RNA replication in a host cell. Preferably, this exogenous RNA segment is located within the open reading frame encoding the STMV coat protein.

The present invention further relates to DNA transcription vectors containing a substantially full-length cDNA copy of such infectious recombinant RNA molecules.

In another aspect, the present invention concerns a method of transforming plant cells, comprising introducing into the cytoplasm of such cells a recombinant RNA molecule derived from STMV ssRNA, having an exogenous RNA segment at a site that is non-essential for RNA replication in such host cells, and either (a) infecting the cell with an helper virus of STMV or (b) otherwise introducing into the cell the genome of an helper virus of STMV. As such, the invention also encompasses a composition which comprises in combination, in a plant cell or in a mixture employed in transforming, or employable to transform, a plant cell, such a recombinant RNA molecule and either an helper virus of STMV or the genome of an helper virus of STMV.

In a still further aspect, the present invention concerns a method for the production of an exogenous protein in the cytoplasm of a plant cell, which has been co-infected with an infectious recombinant STMV RNA molecule as hereinabove described, or with a cDNA having one strand substantially complementary to said RNA and capable of providing said recombinant RNA in vivo, and either (a) an helper virus of STMV or (b) (i) the genome of an helper virus of STMV or (ii) a cDNA having one strand substantially complementary to the genome of such an helper virus and being capable of providing said genome in vivo in the cell, which method comprises

exposing said co-infected plant cell to, or maintaining said co-infected plant cell under, conditions whereby said protein is made in the cell by translation of said infectious recombinant STMV RNA.

5 It is further possible to introduce infectious recombinant STMV RNA or cDNA molecules into plants, following plant transformation with Agrobacterium tumefaciens carrying modified, recombinant Ti plasmids. Since the expression of the exogenous sequences present  
10 in the recombinant STMV RNA molecules is dependent upon infection with an helper virus such as TMV, the helper virus can be used as a regulatory trigger for the expression of the gene(s) foreign to the STMV genome. A particular use for such a system is the introduction of  
15 presumed cross-protection genes from other viruses, such as cucumber mosaic virus, into plants susceptible to TMV, and high-level expression of antisense RNA for control of plant viral diseases. This particular aspect is also within the scope of the invention.

20 The present invention is directed to the above aspects and all associated methods and means for accomplishing such. For example, the invention includes the technology requisite for isolation and purification of STMV ssRNA, transcription procedures, methods for  
25 sequencing genomic RNA, cDNAs and RNA transcripts, plant inoculation techniques, and the like.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts the nucleotide sequence  
30 (genomic sense) of STMV RNA and the deduced amino acid sequences of the two open-reading frames. Stop codons are identified as \*. Regions of the coat protein from which amino acid sequence data were obtained are underlined.

35 Figure 2A is a sequence comparison of the 3'-terminal regions of STMV RNA (middle line of each comparison, bases 770 to 1059), TMV U2/U5 RNA (upper line

of each comparison, 110 terminal 3' nucleotides), and TMV U1 RNA (lower line of each comparison, bases 6100 to 6395). Identity is indicated by a "\*" and gaps generated to provide a best fit are indicated with "----". The stop codon for the coat protein is underlined for TMV U1 and U5.

Figure 2B is a comparison of the 5' sequences of STMV RNA, BMV-RNA3 and CMVQ-RNA3.

Figure 2C is a comparison of the 5' sequences of STMV RNA, STNV RNA and SPMV RNA.

Figure 3 shows the strategy used towards construction of a complete genomic clone for STMV. Clone pSTMV-2 (designated "2" in the Figure) was joined to clone pSTMV-13 (designated "13" in the Figure) at a unique DraI site to produce clone pSTMV2+13. A near full-length clone was created by digestion of double-stranded cDNA of STMV with BglII and HindIII to yield a fragment (designated "31", for clone pSTMV-31, in the Figure) which was ligated to pSTMV2+13 that had been digested with BglII and HindIII. The resulting construct, pSTMV3, shown at the bottom of the Figure, lacks only the 73 5' terminal bases. The heavy line at the top of the figure is a partial restriction map of STMV cDNA.

## DETAILED DESCRIPTION OF THE INVENTION

### Definitions

The term "infectious" used in connection with RNA or cDNA molecules derived from STMV ssRNA indicates that the STMV viruses having such RNA, per se or made from such a cDNA molecule, in their genome are capable of replication and accumulation in living cells. Such molecules do not contain any sequence, either as a result of spontaneous mutation or as a consequence of genetic manipulation, that would disrupt virus RNA replication.

The term "derived from" is used to refer to a viral RNA, typically one that is naturally occurring,

that is modified in a recombinant RNA molecule according to the invention. Recombinant RNAs described herein are derivatives of STMV genomic RNAs. Substantial portions (i.e., segments) of these recombinant RNAs are from STMV genomic RNA. However, components that are endogenous to STMV genomic RNA but are not essential for infectivity may be omitted, or replaced, for example by exogenous RNA segments, in the recombinant RNAs according to the invention that are derived from STMV genomic RNA. In addition, natural mutations or induced changes in the STMV ssRNA segments, which do not eliminate the biological activities of the segments or, if they result in amino acid changes in encoded proteins, do not eliminate the biological activities of the proteins, may be present. For deriving the recombinant RNA molecules according to the present invention from STMV ssRNA, different methods may be employed. The manner of deriving may, for example, be by transcription, direct recombination at the RNA level, or insertion of heterologous DNA segments into cDNA copies of the STMV genomic RNA followed by transcription of the cDNA copies.

The terms "capsid protein" and "coat protein" are used interchangeably and refer to a 17,500 Mr protein encoded by the open reading frame beginning at nucleotide 163 in the authentic nucleotide sequence of STMV ssRNA (see Fig. 1).

The term "substantially full-length copy of an infectious STMV ssRNA" is used to refer to cDNA molecules that may not be exact copies of the respective infectious ssRNAs but comprise the complement of every nucleotide that is required for infectivity. Such cDNA molecules may be shorter or longer than the ssRNA molecules they complement, and even may include nucleotides with no corresponding ribonucleotide in the RNA, as long as the infectivity is preserved.

The term "degenerate equivalent thereof," as used in connection with STMV ssRNA, and RNA of the

invention, or a substantially full-length DNA copy f either, means an equivalent which differs in nucleotide sequence but encodes the same amino acid sequence(s), taking account of the degeneracy of the genetic code.

5 Thus, a multitude of different nucleic acid sequences encoding the same amino acid sequence are considered degenerate equivalents as defined herein. In addition, natural mutations or induced changes in the RNA or DNA segments, which do not eliminate the biological  
10 activities of the segments or, if they result in amino acid changes in encoded proteins, do not eliminate the biological activities of the proteins, fall within the scope hereof and are considered equivalents as well.

The term "exogenous RNA segment," as used  
15 herein, refers to an RNA segment which, in an infectious RNA according to the invention, is heterologous to naturally occurring STMV ssRNA, e.g., a segment which does not occur in such STMV ssRNA. Typically, the exogenous RNA segment will be a segment inserted into the  
20 segment of STMV RNA encoding the coat protein or replacing all or part of such coat-protein-encoding RNA. The exogenous RNA segment will typically itself encode some desired protein. The source of an exogenous RNA segment may be a natural source, including viruses other  
25 than STMV, bacteria, yeasts, fungi, plants and animals; or the exogenous RNA segment may be entirely or partly synthesized. The exogenous RNA fragments may have different functions, including, but not limited to, encoding different heterologous proteins, catalytic  
30 function, regulatory functions. An exogenous RNA segment may also be an anti-sense RNA to a nucleic acid segment that occurs naturally in a target cell, or that may occur in a target cell if the cell is infected by a virus, and whose function is intended to be blocked, interrupted or  
35 otherwise interfered with by hybridization with the recombinant STMV genome of the invention which comprises the exogenous, anti-sense segment.

The amino acids, which occur in the various amino acid sequences referred to in the specification have their usual, three- and one-letter abbreviations, routinely used in the art, i.e.:

5	<u>Amino Acid</u>	<u>Abbreviation</u>	
	L-Alanine	Ala	A
	L-Arginine	Arg	R
	L-Asparagine	Asn	N
	L-Aspartic acid	Asp	D
10	L-Cysteine	Cys	C
	L-Glutamine	Gln	Q
	L-Glutamic Acid	Glu	E
	L-Glycine	Gly	G
	L-Histidine	His	H
15	L-Isoleucine	Ile	I
	L-Leucine	Leu	L
	L-Lysine	Lys	K
	L-Methionine	Met	M
	L-Phenylalanine	Phe	F
20	L-Proline	Pro	P
	L-Serine	Ser	S
	L-Threonine	Thr	T
	L-Tryptophan	Trp	W
	L-Tyrosine	Tyr	Y
25	L-Valine	Val	V

The term "transcription vector" refers to vectors capable of transcribing cDNA sequences contained therein into the corresponding RNA sequences, where such sequences are in operational association with other sequences capable of effecting their transcription, i.e. promoter sequences for an RNA polymerase. In general, transcription vectors usually used in recombinant DNA technology are often in the form of "plasmids", i.e. circular, double-stranded DNA loops which in their vector form, are not bound to the chromosome. In the present specification the terms "vector" and "plasmid" are used



interchangeably. However, the invention is intended to include other forms of transcription vectors as well, which function equivalently.

5 The term "a site that is non-essential for RNA replication" is used to refer to an RNA fragment that is able to tolerate the presence of a non-viral sequence without disrupting RNA replication.

10 The term "co-infection" (by an STMV-originated RNA and a helper virus of STMV) and grammatical variations thereof, mean that two independently replicating virus RNA molecules are caused to occur in a host cell. The term contains no restriction whatsoever, concerning the method and order of infection with the satellite and helper RNA.

15 The transformation of plants cells can be carried out by any method known in the art for introducing RNA or cDNA into plant cells, tissues, protoplasts or whole plants. RNA or cDNA as well as the helper virus, are preferably introduced into whole plants  
20 by mechanical inoculation. Inoculation may be with RNA alone or with virions containing the desired RNA. The selection of the best suited transformation protocol including the manner and parameters of transformation, timing of transformation, etc. is well within the  
25 knowledge of persons of ordinary skill in the art.

The STMV-based transformation/expression system has a broad host range, including many economically important species. The plant hosts include herbaceous plants such as solanaceous plants, e.g. tobacco, tomato,  
30 etc.

The invention is further illustrated by the following non-limiting examples. The techniques utilized in carrying out the invention are well known to those of ordinary skill in the arts of molecular and plant biology  
35 and plant virology. The employed methods are fully described, either directly, or by reference to literature references disclosing their detailed description. Unless

otherwise indicated, the materials used are commercially available, or are in wide circulation among persons practicing in the pertinent arts.

Exogenous gene expression studies were performed with the bacterial chloramphenicol acetyl transferase (CAT) gene, but other genes the expression of which produces an easily detectable product, such as the luciferase gene from firefly, Photinus pyralis, may equally be used in gene expression studies.

#### Example 1

##### Cloning of STMV genome

###### a. Source of STMV

STMV was purified and separated from TMV U5 by rate zonal density gradient centrifugation as described by Valverde and Dodds, J. Gen. Virology **68**, 965 (1987). RNA was isolated from purified STMV by either the sodium perchlorate [Wilcockson and Hull, J. Gen. Virology **23**, 107 (1974)] or the proteinase K procedure [Maniatis et al., "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1982)]. RNAs were denatured in 8 M urea, heated at 60°C for 3 min., and electrophoresed through 1% low gelling temperature agarose (Bethesda Research Laboratories). The intact full-length RNA was recovered from the soft agarose by the freeze/thaw procedure of Benson, BioTech, March/April, 66-67 (1984).

###### b. Direct enzymatic sequencing of STMV RNA 3' and 5' ends and PEI cellulose determination of terminal residues

To determine the complete sequence of the STMV RNA genome, several strategies were utilized. Use of these different strategies yielded sequence data which ruled out possible ambiguities encountered in the use of any single methodology. This was particularly important

in the determination of the 3'- and 5'-terminal sequences of the genome.

The 3'- and 5'-terminal residues of STMV RNA were identified by PEI cellulose (Polygram cell 300 PEI/UV<sub>254</sub>, Brinkmann) thin layer chromatography (Buzayan *et al.*, Virology **151**, 186 (1986)).

Labeling of the 3' terminus of STMV RNA was with [5'-<sup>32</sup>P]pCp [Amersham] using T4 RNA ligase [Donis-Keller *et al.*, Nucleic Acids Res. **4**, 2527 (1977)].

Labeling of the 5' terminus of STMV was with [ $\gamma$ -<sup>32</sup>P]ATP using the forward reaction of T4 polynucleotide kinase (Maniatis, *et al.*, supra).

The 5' terminal nucleotide of kinase labeled STMV RNA was identified by PEI thin layer chromatography. Intact STMV RNA did not require phosphatase treatment to achieve efficient labeling with polynucleotide kinase. Treatment of RNA with TAP and/or BAP prior to using kinase did not increase the efficiency of labeling the 5' termini of STMV RNA. The nucleotide detected after nuclease P1 digestion of treated RNA was predominantly adenosine, indicating that the majority of the 5' termini of STMV RNA encapsidated in virions have a non-phosphorylated adenosine residue.

The 3' terminal nucleotide of T4 RNA ligase-labeled STMV RNA was identified by PEI thin-layer chromatography. The only nucleotide detected after RNase T2 digestion was adenosine.

Direct enzymatic (Donis-Keller *et al.*, supra) sequencing of the 3' end of STMV RNA was conducted as previously described (Zuidema *et al.*, Proc. Natl. Acad. Sci. (USA) **83**, 5019 (1986)). These sequence data were used to construct an oligonucleotide for use in the production of a cDNA library.

#### c. Method of cloning

First strand cDNA synthesis from STMV ssRNA was primed using sonicated and DNase-treated calf thymus DNA

(Maniatis et al., Supra), or by using a deoxyoligonucleotide complementary to the 3' 16 terminal nucleotides of STMV RNA. The sequence of this deoxyoligonucleotide was based on RNase determined sequence data. First strand cDNA synthesis was also primed using oligo dT (12-18) (Maniatis et al., supra) on STMV RNA that had been polyadenylated at its 3' terminus as described by Gething et al., Nature 287,301 (1980).

Second strand synthesis was performed using RNase H and DNA polymerase I (Bethesda Research Laboratories), essentially as described by Gubler and Hoffman, Gene 25, 263 (1983). The resulting double stranded cDNA was either digested with TagI and ligated into the AccI site of pUC118 or pUC119 plasmid (Vieira and Messing, Production of Single-Stranded Plasmid DNA, In "Methods in Enzymology" (R. Wu and L. Grossman, Eds. 153, 3-11 Academic Press, New York (1987)), or was made blunt-ended with T4 DNA polymerase as described by Gubler and Hoffman, supra, and ligated into the SmaI site of plasmid pUC118 or pUC119. Ligation products were used to transform E. coli strain DH5 $\alpha$  (Bethesda Research Laboratories).

Ampicillin resistant transformants were selected by colony hybridization with kinase-labeled STMV virion RNA probe (Rezaian et al., Virology 131 221 (1983)). Nick-translated DNA from four of the largest plasmids averaging 450 base pairs hybridized only to STMV RNA in Northern hybridizations, which confirmed insert specificity.

d. Sequencing of cDNA clones to determine the sequence of the STMV genome

Recombinant plasmids were purified from 29 ampicillin-resistant colonies that hybridized with kinase-labeled STMV RNA probe. Their inserts were excised by digestion with TagI, or KpnI and PstI and resolved by electrophoresis in agarose gels. Sequencing

was by the dideoxynucleotide chain termination method (Sanger et al., Proc. Natl. Acad. Sci. (USA) 74 5463 (1977)). Certain of the inserts were sequenced in both directions. When ambiguities arose and there was a need to confirm sequence data, genomic RNA itself was sequenced using deoxyoligonucleotides as primers. Thus, from the 3'-terminus to about ribonucleotide 830 and from about nucleotide 470 to the 5'-terminus of the genomic RNA were sequenced directly. The 29 cDNA fragments, including those in clones pSTMV-2, pSTMV-13, and pSTMV-31 (Fig. 3), represented the entire genome of STMV with the exception of the nine 5'-terminal nucleotides.

To complete the 5' sequence data of STMV RNA, a deoxyoligonucleotide (5'-GGCGACTGAAGGCC) complementary to nucleotides 102-115 was used to prime reverse transcription in the presence of dideoxynucleoside triphosphates (Palmenberg et al., Nucl. Acids Res. 12, 2969 (1984)). The sequence of the 5' terminus of STMV RNA thus was deduced from gels in the standard fashion. RNase determined sequence was used for confirmation. RNase determined sequence was also used to confirm that cDNA clones covered the the 3' terminus.

The primary structure of STMV RNA, shown in Fig. 1, is 1,059 nucleotides long. Restriction maps of six of the larger cDNAs used to determine the sequence in Fig. 1 were obtained using 30 different restriction enzymes; these maps conformed exactly to the maps predicted from the nucleotide sequence in Fig. 1.

## Example 2

### a. Sequencing of clone inserts

The STMV RNA sequence was screened for potential coding regions in all three reading frames in both the virion (positive) and complementary (negative) senses. Two open reading frames (ORFs) with the potential to code for proteins of 6,800 Mr and 17,500 Mr were observed. The first ORF (Fig. 1) begins at the

first 5' AUG triplet in positions 53 to 55, and ends with a UGA termination codon at bases 227 to 229. The largest ORF (Fig. 1) begins at the second AUG triplet in positions 163 to 165 and ends with a UGA termination codon at residues 640 to 642. The 418 nucleotides following the ORF for the capsid protein do not encode any polypeptides longer than 16 amino acids. The negative sense RNA has a single ORF which had the coding capacity for an 8K protein.

The 12 N-terminal amino acids of the capsid protein, corresponding to the first 12 codons of ORF 2, are underlined in Fig. 1. There is a perfect fit between the predicted amino acids and those determined by amino acid sequencing from the N-terminus. The results of an amino acid composition analysis agreed well with the empirically determined amino acid composition of the STMV capsid protein indicated in Fig. 1.

#### b. Characterization of sequence

One microgram of STMV RNA was used in each translation reaction in either a rabbit reticulocyte lysate or a wheat germ extract cell-free system (Promega Biotec, Madison, WI, U.S.A.) using <sup>35</sup>S labeled methionine. Time course analysis of cell-free translation products was achieved by removing one microliter aliquots from a single reaction mixture at 1 min. intervals. <sup>35</sup>S labeled polypeptides were denatured by boiling in 63 mM Tris-Cl, pH 6.8, 2% SDS, 10% glycerol, and 5% mercaptoethanol (denaturing buffer), then electrophoresed through 15% polyacrylamide, 0.1% SDS gels (Laemmli, Nature 227, 680, (1970)). Gels were soaked in three changes of 30% methanol/10% acetic acid, vacuum dried at 60°C, and exposed to Kodak X-Omat AR film for 24 to 72 hr. STMV RNA efficiently directed the synthesis of two polypeptides in both wheat germ and rabbit reticulocyte systems.

Samples for Western-blot analysis included cell-free translation products, STMV virions, TMV U5 virions, homogenized tobacco tissue infected with STMV/TMV U5, and homogenized non-infected tobacco tissue. Samples were treated and electrophoresed as described above. After electrophoresis, proteins were electrophoretically transferred to a nitrocellulose membrane (Trans-Blot Cell, Bio-Rad). After immobilization of the proteins to the membrane, the blot was developed as described by Blake et al., Anal. Biochem. 135, 175 (1984) using a 1/1000 dilution of an STMV specific antiserum (Valverde and Dodds, supra).

Western-blot analysis of these in vitro translation products revealed that the larger protein had the same electrophoretic mobility as the STMV capsid protein (17,500 Mr) and an equivalent protein found in sap from STMV infected plants. These proteins reacted with an antiserum specific for STMV. The second protein (6,800 Mr) and TMV capsid protein did not react with STMV antiserum. The 6,800 Mr protein corresponds in size to the putative protein coded by the first ORF in the genome. Time course analysis of in vitro translation demonstrated that the ratio between the amounts of the 6,800 Mr protein and the 17,500 Mr protein were similar at each sampling time. No significant amount of higher molecular weight proteins were synthesized at any sampling time.

It can be concluded from size analysis of cell-free translation products, serological analysis, amino acid sequence and composition results that ORF 2 encodes the 17,500 Mr STMV capsid protein.

### Example 3

#### a. Construction of pSTMV+2Δ6

Methods for the synthesis of cDNA clones were as herein above described. Freshly prepared STMV ssRNA purified from STMV virions was obtained from J. A. Dodds,

University of California, Riverside. Clones pSTMV-2, pSTMV-13, and pSTMV-31 were constructed at the University of California, Riverside, following Example 1 and obtained from J. A. Dodds. As shown in Figure 3, 100  
5 nanograms of insert DNA fragments from these clones were ligated at a unique DraI and unique BglII site. This ligation mixture was transformed into E. coli strain DH5 $\alpha$ , and transformants were selected on ampicillin plates. Plasmid DNA isolated from these transformants  
10 was digested with KpnI and HindIII to identify near full-length STMV cDNA clones. One of these clones, pSTMV-3, which contains a 986 base pair insert which lacks the 73 5' terminal nucleotides of STMV, was sequenced to confirm identity with the STMV consensus sequence and then used  
15 for the construction of a full-length STMV cDNA clone.

To generate 5'-terminal clones, an oligonucleotide complementary to nucleotides 351-362 (5'-GGAATCTGTCCGG) of the STMV genomic RNA was used to prime first-strand cDNA synthesis. A second primer  
20 (5'-GGTACCAGTAAACTTACCAATCAAAA) complementary to the 3' end of the first-strand cDNA was then used to specifically prime second strand cDNA synthesis of only those molecules that extend through (are complementary to) the final 5'-terminal nucleotides of the STMV genome.  
25 Included in this oligonucleotide primer was a 6-base overhang that is the recognition sequence for KpnI. This unique restriction site was used to facilitate cloning of the 5'-end of the STMV genome without losing 5'-terminal sequences, which is a deficiency of many alternative  
30 cloning strategies.

Approximately 5  $\mu$ g of double-stranded cDNA derived from the above-described reactions was digested with KpnI and MstII to yield a 249 base pair fragment and a 113 base pair fragment. These fragments were not  
35 separated as only the 249 base pair fragment contained both the KpnI 3' overhang and MstII 5' overhang. pSTMV-3 DNA was digested with KpnI and MstII; the resulting



vector fragment was gel purified, and .50 nanograms was ligated with ~ 100 nanograms of the 249 base pair KpnI-MstII cDNA fragment. This ligation mixture was used to transform E. coli strain DH5 $\alpha$ , and transformants were  
5 selected on ampicillin plates. Transformants containing 5' terminal sequences were identified by colony hybridization with an oligonucleotide  
5'-GCATTAATGACCACGACAGTCCTGGTGGTTAGGTCTTTTGATTGGTAAGTTT  
TATTGGTACC) complementary to the 5' end of genomic STMV  
10 RNA. Clones thus identified were anticipated to contain full-length copies of the STMV genome.

Sequence analysis (Sanger dideoxynucleotide method on denatured plasmid DNA using <sup>35</sup>S-dATP) indicated that all nine clones derived from these experiments fell  
15 short of the 5' terminus by nine nucleotides. The reasons for this were unclear, although sequence analysis indicated that there was a tetranucleotide sequence within the STMV genome, i.e., bases nine to twelve (5'-TTAC), at which KpnI (recognition sequence 5'-GGTACC)  
20 may have digested the double-stranded cDNAs. New England Biolabs has recently reported apparent star activity in preparations of KpnI, supporting the hypothesis that KpnI could cut these clones at this site. One of these clones, pSTMV-KMB35, was used in further constructions  
25 designed to yield a full-length STMV clone.

To produce insert DNAs containing the missing nine 5'-terminal nucleotides, an alternative cloning strategy was used. First-strand cDNA synthesis was primed using an oligonucleotide (5'-CCCTTCGATTTAAG)  
30 complementary to nucleotides 940 through 958. This was chosen such that a unique restriction enzyme site (BglII) could be utilized for forced orientation cloning using pSTMV-KMB35 as a vector, as sequences 3' of this site are difficult to clone due to secondary structure. Second  
35 strand synthesis was conducted using an oligonucleotide to prime second strand synthesis of only molecules that extend to the 5' terminus. This oligonucleotide

(5'-CTGCAGATCTAGAATAAACTTACCAATCAAAAG) utilized a unique XbaI extension to overcome the problem discussed above which resulted in clones that were missing the nine 5'-terminal nucleotides. This restriction site (XbaI) also provided the basis for a convenient method for subcloning insert DNA into the pMJ5 transcription vector (described in detail in Janda *et al.*, Virology 158, 259 (1987)) in such a way that only two non-viral nucleotides are included in transcripts derived from this construct.

Approximately 5  $\mu$ g of this double-stranded cDNA was digested with XbaI and BglII to yield a 791 and a 149 base pair fragment. The 791 base pair fragment was not further purified because only it contained both the XbaI 5' overhang and the BglII 5' overhang. pSTMV-KMB35 DNA was digested with XbaI and BglII and gel purified. Approximately 100 nanograms of the 791 base pair XbaI-BglII cDNA fragment was ligated with ~ 50 nanograms of the vector fragment. This ligation mixture was transformed into *E. coli* strain DH5 $\alpha$ , and transformants were selected on ampicillin plates. Plasmid DNA was prepared from a transformant (pXBH118A); when digested with XbaI and HindIII, an insert fragment of ~ 1060 base pairs was produced. This insert DNA was sequenced by the Sanger dideoxynucleotide method on denatured plasmid DNA using <sup>35</sup>S-dATP, and the sequence data revealed that the missing nine 5'-terminal nucleotides had been cloned in clone pXBH118A. However, three transitions had occurred in the 3' non-coding region of the genome: C to U at position 682, A to C at position 751, and C to U at position 753. Insert DNA derived from this clone (pXBH118A) was utilized in construction of an RNA transcription vector pSTMV+2A6 as described below.

Five  $\mu$ g of pMJ5 DNA was digested with StuI and HindIII, and the resultant vector DNA fragment was gel purified. Insert DNA was prepared from pXBH118A DNA by digestion with XbaI followed by treatment with mung bean

nuclease to remove the non-viral nucleotides. This DNA preparation was then digested with HindIII. After gel purification, ~ 100 nanograms of the resultant 1059 base pair fragment was ligated with 100 nanograms of the prepared pMJ5 vector DNA. This ligation mixture was transformed into E. coli strain DH5 $\alpha$ , and transformants were selected on ampicillin plates. Plasmid DNA purified from these transformants was digested with KpnI and HindIII to identify those containing an insert fragment of ~ 1120 base pairs. The insert DNA from one clone, pSTMV+2 $\Delta$ 6, was sequenced by the Sanger dideoxynucleotide method on denatured plasmid DNA using <sup>35</sup>S-dATP to confirm that this clone contained the correct sequence. It was determined from the sequence data that the insert DNA contained three errors in addition to the three transitions that are described above. To reflect the presence of these six errors, this vector was designated pSTMV+2 $\Delta$ 6. To correct all six errors, another vector, pSTMV+2 $\Delta$ 1, otherwise referred to herein also as pSTMV+2 $\Delta$ 2, was constructed.

b. Construction of pSTMV+2 $\Delta$ 1

Approximately five  $\mu$ g of pSTMV+2 $\Delta$ 6 DNA was digested with SalI and HindIII, and the fragments were separated by agarose gel electrophoresis. The ~ 4300 base pair fragment comprising the vector and 5' end of the STMV genome was purified.

An oligonucleotide

(5'-GATTTAAAGCTTGGGCGCTTACCCGCGGTTAGGG-3') complementary to the correct 3' terminal sequence of the STMV genome and containing a unique HindIII extension was synthesized. This oligonucleotide was used to prime first-strand synthesis of cDNA. The resultant cDNA molecules were made double-stranded, digested with SalI and HindIII, and the 450 base pair SalI-HindIII fragment was purified from an agarose gel. One hundred nanograms

of this fragment was ligated with fifty nanograms of the HindIII-SalI fragment derived from pSTMV+2Δ6. This ligation mixture was transformed into E. coli strain DH5α, and transformants were digested with KpnI and HindIII to identify clones containing an ~ 1120 base pair insert fragment. The insert DNA was sequenced by the Sanger dideoxynucleotide method on denatured plasmid DNA using <sup>35</sup>S-dATP. This sequence analysis indicated that all six errors present in pSTMV+2Δ6 had been corrected.

However, two new errors (a C to G transition at nucleotide 1057 and a C to T transition at nucleotide 984) had been introduced, and the resultant vector was first designated pSTMV+2Δ1, later changed to pSTMV+2Δ2, to reflect these two nucleotide differences from the STMV RNA genome.

c. Construction of pSTMV+2ΔOP

A 3' PstI site was introduced at the 3' end of the STMV sequence through use of the oligonucleotide:

5'-ATCTGCAGGGCCGCTTACCCGCGGTTAGGG-3'. This oligonucleotide is complementary to the 3' terminal 24 STMV nt sequence and additionally contains a PstI recognition site. The oligo was used to prime first strand cDNA synthesis from STMV RNA. Second-strand was synthesized by the nick-translation method using RNase H and DNA polymerase I (Gubler and Hoffman, 1983). The resultant double-stranded cDNAs were made blunt-ended with T4 DNA polymerase and digested with PstI. 50 ng of insert were ligated to 100 ng of SmaI-PstI-digested pUC118 and the ligation was transformed into DH5-α cells and amp<sup>R</sup> colonies were selected. Correct plasmid was first identified by bands which resulted from digestion with PstI and BglII, and then confirmed by DNA sequencing. One plasmid that contained the PstI site and the correct 3' terminal 500 nt was called pSTMVPstOligo#6.

Plasmid pSTMVPstOligo#6 was digested with EcoRI and BglII and the ~3390bp fragment was isolated on a 1.0% agarose gel. 50 ng of this fragment were ligated to 100 ng of the 833 bp EcoRI-BglII fragment of pSTMV+2Δ2 isolated from a 1% agarose gel. The ligation was transformed into DH5-alpha cells and amp<sup>R</sup> colonies were selected. Correct plasmid was identified by sequencing, and was called pSTMV+2ΔOP. The insert in this plasmid has the exact sequence of STMV, as determined in Example 1.

d. Construction of pSTMV+2ΔHPA

Plasmid pSTMV+2Δ2 was digested with HpaI and the large vector fragment was isolated on a 1.0% agarose gel. The ends of the isolated vector fragment were religated and the ligation mixture was transformed into DH5-alpha cells. Amp<sup>R</sup> colonies were selected. Correct plasmid exhibited one band of 3850 bp upon digestion with HpaI, and was called pSTMV+2ΔHPA. This plasmid contains a deletion from nucleotides 414 through 787 of the STMV coat protein and 3' noncoding sequence.

e. Analysis of RNA transcripts

pSTMV+2Δ6, pSTMV+2Δ2, pSTMV+2ΔOP, and pSTMV+2ΔHPA were used to produce transcripts that have only two non-viral nucleotides at the 5' terminus, in addition to six, two or zero transitions at the 3'-end, or a deletion at the 3'-end, respectively. Forty μg of vector DNA were linearized with HindIII (2Δ6, 2Δ2, 2ΔHPA), and then treated with 20 units of mung bean nuclease (New England Biolabs) to remove the four non-viral nucleotides contained in the HindIII 5' overhang, or linearized with PstI (2Δ6), and then, in all four cases, phenol\chloroform extracted, and ethanol precipitated. Ten μg of these DNAs were incubated in the presence of 100 units of T7 RNA polymerase in a total volume of 100 μl containing 40 mM Tris-HCl, pH 7.5, 6 mM

MgCl<sub>2</sub>, 2 mM spermidine, 10 mM NaCl, 10 mM DTT, 40 units RNasin, and 2.5 mM each ATP, CTP, UTP, and GTP.

Reactions were incubated at 37°C for two hours and typically yielded 60-70 µg of transcribed RNA. The

5 transcripts resulting from pSTMV+2Δ6 and pSTMV+2Δ2 migrated with the same mobility as did wild type STMV RNA on denaturing agarose gels. The transcripts from STMV+2ΔOP similarly migrated with the same mobility, while those from pSTMV+2ΔHPA were smaller than wild-type  
10 STMV.

f. In vitro translation

Transcripts derived from pSTMV+2Δ6 and pSTMV+2Δ2 translated efficiently in a wheat germ cell-  
15 free extract translation system (Promega Biotec) and produced polypeptides identical in size and relative amounts to those produced from wild type STMV ssRNA.

Example 4

20 Infectivity studies

a. Infectivity protocol

For infectivity studies, 0.9 ml of a mixture containing 0.05 M glycine, 0.03 M K<sub>2</sub>HPO<sub>4</sub>, pH 9.2, 1% sodium pyrophosphate, 1% Macaloid, and 1% Celite,  
25 adjusted to pH 8.5 with H<sub>3</sub>PO<sub>4</sub>, was added to a 100 µl transcription reaction mixture. One hundred µl of this 1000 µl mixture was utilized for each mechanically inoculated tobacco (N. tabacum cv. Xanthi) leaf. Inoculations were on leaves systemically infected with  
30 the TMV U5 or U2 helper virus. Alternatively, leaves were inoculated with helper virus 0, 1, 2 or 6 days before inoculation with transcript RNA.

A mixture of two different tobamoviruses were used in this study, TMV U2 and TMV U5. TMV U5 is the  
35 natural helper virus for STMV. TMV U5 was obtained from J.A. Dodds, University of California, Riverside, and was derived from a stock of Nicotiana Tabacum cv. Xanthi

originally developed and maintained by L.G. Wheathers. TMV U2 is available from the American Type Culture Collection, Rockville, Maryland, USA and is essentially the same as TMV U5, merely having been isolated first at a geographical location different from that at which TMV U5 was first isolated. The isolates used had been passaged through single local lesions in Nicotiana tabacum cv Xanthi nc. prior to inoculation to N.tabacum cv Xanthi, which was the host used for maintaining isolates. Stock plants harboring the viruses have been tested repeatedly for the presence of STMV and have always tested negative. However, prior to each use as a source of inoculum, extracts from the maintenance hosts were tested for STMV contamination by dot spot hybridization or immunodiffusion assay. All plants indexed negatively throughout the experimental period.

TMV U5-infected Nicotiana tabacum cv. Xanthi are being maintained as stock plants under standard conditions (Valverde and Dodds, Supra and J. Gen. Virology 67, 1875 (1986)). One of the two TMV U5 isolates used was determined by dsRNA analysis (detection of a pair of RF dsRNAs with slightly different mobilities and other pairs of dsRNAs with strain-specific mobilities) and host reaction in N. sylvestris (local lesion formation indicative of TMV U1) to be contaminated by readily detectable levels of TMV U1 and is referred to as TMV U5B to distinguish it from TMV U5A which was not detectably contaminated with TMV U1 by these criteria.

Freshly prepared STMV ssRNA, purified from STMV virions, was obtained from J.A. Dodds, University of California, Riverside.

Seeds of N. tabacum cv. Xanthi were provided by J.A. Dodds, University of California, Riverside. Plants were grown and maintained under standard conditions.

#### b. STMV Assays

##### 1. Immunoanalysis for STMV coat protein

Ten to 21 days post-inoculation, the uppermost 2-3 non-inoculated systemically infected leaves (1-2 g fresh weight) were harvested for analysis. Ouchterlony double-immunodiffusion assays were carried out to screen for the presence of STMV coat protein antigen and to look for serological differences between a wild type STMV isolate in a mixed infection with TMV U5, and progeny virions from RNA transcript infections. One g of leaf tissue was ground in 1 ml of 0.14M NaCl, the extract was filtered through Miracloth and 50  $\mu$ l of the sample was placed in alternating wells (wild type sample next to experimental sample) cut in a 1.0% agarose, 0.1M NaCl plate. An STMV specific antiserum was used in the center well (Valverde and Dodds, 1987, supra). Initial results were recorded after 24 hr. incubation at room temperature; examination for spur formation between adjacent samples was stopped six days later.

## 2. Nucleic acid analysis of STMV sequence

Total nucleic acids were extracted as described by Dunsmuir et al., and spotted onto Zeta-Probe membrane (BioRad). Alternatively, 20  $\mu$ l of each test sample used for the Ouchterlony assay were placed in a microfuge tube and centrifuged for 6 minutes. Two to five  $\mu$ l were applied to a nitrocellulose membrane, air dried and baked at 80°C for 2 hours. Blots were assayed with <sup>32</sup>P-labelled plasmid pSTMV2, which contains a cDNA insert that represented nucleotides 70 to 405 of the STMV genome, or with RNA transcript from similar plasmids, or cDNA transcribed from STMV genomic RNA. Extracts from non-inoculated plants or plants infected with any helper virus alone did not give a positive signal with any of the probes.

## 3. Analysis of STMV virions

Wild-type STMV or transcript-derived virions were purified and separated from TMV U5 and U2 by rate



zonal density gradient centrifugation as described by Valverde and Dodds (1987). Relative banding positions were compared between wild-type and transcript-derived virions. Samples of gradient-purified virions were  
5 negatively stained with 2% uranyl acetate and examined by electron microscopy. RNA was isolated as described in Example 1.a, supra.

c. Analysis of results of infectivity studies

10 1. Preliminary study with pSTMV+2Δ6 and pSTMV+2Δ2

At 15 days or 21 days post-inoculation, leaf tissue was harvested from infected plants, and total nucleic acid was prepared according to the first three  
15 steps of the procedure described in Dunsmuir et al., Plant Molecular Biology Manual C1: 1-17 (1988). These nucleic acids were then applied to nitrocellulose filters for dot blot hybridization assays using nick-translated pSTMV2 as a probe. Transcripts derived from pSTMV+2Δ6  
20 yielded three infected plants out of a total of eight plants inoculated with transcripts. Transcripts derived from pSTMV+2Δ2 yielded 100% infectivity.

To rule out the possibility that a wild type STMV contaminant was responsible for the production of  
25 virus in plants inoculated with in vitro transcripts, virions were purified from plants which had been inoculated with transcripts derived from pSTMV+2Δ6. RNA was purified from these virions, and sequence data were obtained. These data included a region of the genome in  
30 which pSTMV+2Δ6 differs at three positions from the wild type. As expected, the C to T transition at nucleotide position 681, the A to C transition at nucleotide position 751, and the C to T transition at nucleotide position 753 all were retained in the virion RNA prepared  
35 from plants inoculated with in vitro transcripts. These results indicated the in vitro transcripts, and not a wild type contaminant, were responsible for the

production of virus in these plants. This result also implied that at least some changes in nucleotide sequence could be accommodated by the virus without loss of infectivity. Whether such changes could include additional sequences and/or deletions was explored in the development of recombinant STMV vectors.

## 2. Four further experiments

### a. Infectivity

Data were obtained from several subsequent experiments involving coinfection of plants with pSTMV+2Δ2 and TMV U2 as helper followed by analysis of leaves 10 days after inoculation. STMV in vitro transcripts that had been stored frozen for up to 12 weeks retained high levels of infectivity. 100% infection was obtained with concentrations from 5 to 50 μg/ml. Incidence of infection by STMV dropped to 75% at 0.5 μg/ml and was not detected when 0.1 μg/ml was used.

Further experiments were carried out with analysis of leaf tissue between 10 and 21 days post-inoculation with STMV transcript at 50 μg/ml, used either without storage or after storage for four weeks. It was possible to obtain 100% infectivity of plants by STMV in most experiments when TMV-U5-A or TMV U2 were used as helper viruses and when pSTMV+2Δ2 and pSTMV+2ΔOP were used to produce transcripts. Delay of 0 hours or 48 hours between inoculation with TMV and inoculation with STMV transcript appeared to have no effect on infectivity. When TMV U5-B was used with these transcripts, the incidence of infection ranged from 25% to 100% depending on the experiment. The lowest incidences of infection (25-50%) were obtained when TMV U5B was the helper and pSTMV+2Δ6 was used to produce transcripts; the percentages were near 100 % with pSTMV+2ΔOP and pSTMV+2Δ2.

CDNA constructs also were prepared to produce (-) sense in vitro transcripts. These transcripts were

not infectious when inoculated under the conditions described above.

When large plants were pre-infected for 48 hours with TMV before inoculation with in vitro transcripts from pSTMV+2Δ2, it was possible to approach 100% infectivity with the STMV construct. The same result was achieved with small plants. No difference in the incidence of infection was seen when helper virus and STMV RNA transcripts were inoculated together or when there was a 24 or 48 hour delay after TMV inoculation before inoculation with the transcripts. Shading plants prior to inoculation provided no advantage, nor did it reduce the incidence of infection.

b. Analysis of progeny virions and RNAs  
Virions indistinguishable from STMV accumulated in plants infected with in vitro transcripts from pSTMV+2ΔOP, pSTMV+2Δ2, and pSTMV+2Δ6, as determined by electron microscopy of negatively-stained particles, analytical sucrose density gradient centrifugation, and Ouchterlony immunodiffusion assay. Similar amounts (4.6 to 4.8 mg/10 g fresh leaf tissue) of STMV virions were purified from plants infected with either STMV or with the three different transcripts. Two major translation products (17.5 Kd and 6.8 Kd) were detected when STMV RNA, transcripts from pSTMV+2Δ2, and progeny RNA from virions derived from pSTMV+2Δ2 transcript infections were used to direct synthesis of peptides in a wheat germ cell-free translation system.

Primer extension in the presence of dideoxynucleotides and RNase sequence analysis were performed to determine if any of the eight different single nucleotide transitions were maintained in the progeny from plants inoculated with STMV in vitro transcripts. RNase sequence analysis of progeny RNA derived by infection with in vitro transcripts prepared from pSTMV+2Δ2 showed that transitions at positions 989 and 1057 were not maintained and that the 3'terminal

sequence of progeny was identical to that of wild-type. RNase sequence analysis of progeny derived infection with in vitro transcripts prepared from pSTMV+2Δ6 gave a similar result in that the transitions introduced at positions 1053 and 1046, and the insertion at position 1055 had reverted to wild-type sequence. However, a single deletion not present in the transcript was detected at residue 1050 causing the sequence to differ from wild-type. Primer extension analysis of this progeny RNA in the region from positions 682 to 753 showed that the three transitions in this region were retained.

The 5' termini of STMV, in vitro transcripts, and progeny derived by infection with these transcripts were examined to determine if the non-viral Gs (guanylates) present in the transcripts were retained in the progeny. Primer extension in the presence of dideoxynucleotides indicated that the progeny RNA was the same length as wild-type STMV RNA and that the transcripts were at least two nucleotides longer than wild-type RNA.

#### Example 5

##### Construction of recombinant STMV vectors

- a. Recombinant STMV-CAT vector system
  1. pSTMV+2Δ6/CAT

The bacterial chloramphenicol acetyl transferase (CAT) gene was inserted into vector pSTMV+2Δ6 near the 3' end of the coat protein coding region into the AccI site at nucleotide position 604. The CAT gene was excised from plasmid pCamVCN (Pharmacia) as a 788 base pair TaqI fragment and gel purified. One hundred nanograms of this fragment was ligated to 50 nanograms of AccI-digested pSTMV+2Δ6 DNA that had been treated with calf alkaline intestinal phosphatase. The ligation mixture was transformed into DH5α bacterial cells, and ampicillin resistant colonies were selected. Plasmids

containing the CAT gene in the same orientation as the viral coat protein gene yielded two fragments (4155 and 892 base pairs) upon digestion with EcoRI. The junction of these two genes results in an out-of-frame fusion protein; hence, the CAT gene is expressed as a result of plant ribosomes initiating protein synthesis from the bacterial CAT initiation codon. This construct was designated pSTMV+2Δ6ACC::CAT(+).

2. Construction of pSTMV+2Δ2/CAT  
expression vector  
(pSTMV+2Δ2ACC::CAT(+))

The CAT gene was inserted into the STMV sequence near the 3' end of the coat protein-encoding region in vector pSTMV+2Δ2.

Plasmid pCamVCN was digested with TagI and the 788 bp CAT-encoding fragment was isolated on a 1.0% agarose gel. One hundred nanograms of the 788 bp fragment were ligated to 50 ng of AccI-digested, calf intestinal phosphatase-treated pSTMV+2Δ2. The ligation was transformed into DH5-alpha cells and amp<sup>R</sup> colonies were selected. Correct plasmid yielded fragments of 850 and 4159 bp upon digestion with EcoRI, and was called pSTMV+2Δ2ACC::CAT(+). The CAT gene is out of frame in relation to the STMV coat protein coding sequence.

3. Construction of pSTMV+2ΔOP/CAT  
expression vector  
(pSTMV+2ΔOPACC::CAT(+))

Plasmid pSTMV+2Δ2ACC::CAT(+) was digested with KpnI and BglII and the 3385 bp fragment was isolated on a 1% agarose gel. 100 ng of the 1620 bp KpnI-BglII fragment from 2Δ2ACC::CAT(+) were ligated to 50 ng of the 3385 bp fragment. The ligation mixture was transformed into DH5-alpha cells, and amp<sup>R</sup> colonies were selected. Correct plasmid demonstrated bands of 850 and 4159 bp upon digestion with EcoRI, and was called pSTMV+2ΔOPACC::CAT(+). The CAT gene is in the proper

orientation, but is out of frame in relation to the STMV coat protein coding sequence in this plasmid.

#### 4. Infectivity studies

In vitro transcripts were prepared from  
5 pSTMV+2Δ6ACC::CAT(+), pSTMV+2Δ2ACC::CAT(+), and  
pSTMV+2ΔOPACC::CAT(+) as described previously and used to  
inoculate tobacco plants.

Analysis of infected and control plants three  
and six days post inoculation showed that CAT was  
10 expressed in the plants infected with  
pSTMV+2Δ6ACC::CAT(+) and was not expressed in plants  
infected with pSTMV+2Δ6. CAT assays were conducted  
according to Herrera-Estrella et al., Plant Molecular  
Biology Manual B1, S.B. Gelvin and R.A. Shilperoort,  
15 Eds., Kluwer Academic Publishers, Dordrecht, The  
Netherlands, pp 1-22 (1988).

CAT protein was not detected in plants infected  
with pSTMV+2Δ2ACC::CAT(+) and helper U5, however dot/spot  
analysis performed 21 days post inoculation revealed the  
20 presence of STMV and CAT nucleic acid in the leaves of  
one of four infected plants. Thus, pSTMV+2Δ2ACC::CAT(+)  
is weakly infective in combination with helper U5.

One-hundred percent of the leaves and systemic  
tissue analyzed from plants infected with  
25 pSTMV+2ΔOPACC::CAT and U2 helper revealed the presence of  
CAT and STMV nucleic acid. Because the CAT gene is out  
of frame in this construct, CAT protein was not expected  
necessarily to be expressed and was not detected.

30 While the present invention has been described  
with some particularity herein, those of skill in the  
pertinent arts will recognize many variations and  
modifications of what has been described that are within  
the spirit of the invention. It is intended that such  
35 variations and modifications are also encompassed by the  
description and following claims.

## CLAIMS:

1. An infectious recombinant RNA molecule derived from STMV ssRNA.
2. An infectious recombinant RNA molecule according to Claim 1 unaccompanied by endogenous components not essential for its infectious properties.
3. An infectious recombinant RNA molecule according to Claim 1, which is made by in vitro transcription of a cDNA.
4. An infectious recombinant RNA molecule according to Claim 3, which is derived from an STMV ssRNA having a nucleotide sequence as shown in Figure 1 or a degenerate equivalent thereof.
5. An infectious recombinant RNA molecule according to Claim 4, which is derived from an STMV ssRNA having the nucleotide sequence as shown in Figure 1.
6. A DNA transcription vector which comprises a substantially full-length cDNA copy of an infectious STMV ssRNA and from which said cDNA is capable of being transcribed.
7. A DNA transcription vector according to Claim 6 wherein the transcript of said substantially full-length cDNA copy of an STMV ssRNA has a nucleotide sequence as shown in Figure 1 or a degenerate equivalent thereof.
8. A DNA transcription vector according to Claim 7 wherein the transcript of said substantially full-length cDNA copy of an STMV ssRNA has the nucleotide sequence shown in Figure 1.
9. A DNA transcription vector according to any one of Claims 6 to 8 wherein transcription of said cDNA is under the control of a promoter for a DNA-dependent RNA polymerase.
10. A DNA transcription vector according to Claim 9, which is a derivative of the plasmid pMJ5 wherein said cDNA is ligated between the StuI and HindIII sites of plasmid pMJ5s.

11. A vector according to Claim 10 selected from the group consisting of pSTMV+2Δ6, pSTMV+2Δ2, and pSTMV+2ΔOP.

5 12. An infectious recombinant RNA molecule derived from STMV ssRNA, having an exogenous RNA segment at a site that is non-essential for RNA replication in a host cell.

10 13. An infectious recombinant RNA molecule according to Claim 12, wherein said exogenous RNA segment encodes a protein.

14. An infectious recombinant RNA molecule according to Claim 12, wherein said exogenous RNA segment comprises RNA having regulatory or catalytic properties.

15 15. An infectious recombinant RNA molecule according to Claim 13 wherein said exogenous RNA segment is derived from a plant virus.

16. An infectious recombinant RNA molecule according to Claim 14 wherein said exogenous RNA segment is derived from a plant virus.

20 17. An infectious recombinant RNA molecule according to Claim 13 wherein said exogenous RNA segment is derived from an animal virus.

25 18. An infectious recombinant RNA molecule according to Claim 14 wherein said exogenous RNA segment is derived from an animal virus.

19. An infectious recombinant RNA molecule according to any one of Claims 12 to 18, wherein said exogenous RNA segment is located within the region encoding the coat protein of STMV.

30 20. A DNA transcription vector containing a substantially full-length cDNA copy of a recombinant RNA molecule according to any one of Claims 12 to 19.

35 21. A DNA transcription vector according to Claim 20, selected from the group consisting of pSTMV+2Δ6ACC::CAT (+), pSTMV+2Δ2ACC::CAT, and pSTMV+2ΔOPACC::CAT.



22. A method of transforming a plant cell, comprising introducing into the cytoplasm of said cell an infectious recombinant RNA molecule according to any one of Claims 13 or 14 and a helper virus of STMV.

5 23. A method according to Claim 22 wherein said helper virus is a strain of TMV selected from the group consisting of U1, U2 and U5.

24. A method according to Claim 22 wherein said RNA molecule and said helper virus are introduced into the  
10 cytoplasm of said cell by mechanical inoculation.

25. A method of Claim 22, wherein said cell is that of a herbaceous plant.

26. A method of Claim 25 wherein said cell is that of a solanaceous plant.

15 27. A method of transforming a plant cell, comprising introducing into the cytoplasm of said cell an infectious, substantially full-length cDNA copy of an STMV ssRNA.

28. A method for the production of an exogenous  
20 protein in the cytoplasm of a plant cell, comprising co-infecting said plant cell with an infectious recombinant RNA molecule according to Claim 13 or with a cDNA having one strand substantially complementary to said RNA and capable of providing said recombinant RNA by  
25 transcription in vivo, and a helper virus of STMV.

29. A method according to Claim 28, wherein said helper virus is a TMV of a strain selected from the group consisting of U1, U2 and U5.

/

```

1  AGUAAAACUUACCAAUAAAAAGACCUAACCAACAGGACUGUCGUGGUCAUUUAUGCUGUU  60
                                     M L L
61  GGGGGACAUAGGGGAAAAACAUUUGCCUUCUUCUACAAGAGGCCUUCAGUCGCCAUAAU  120
    G D I G G K H I A F F Y K R P S V A I I
121  UACUUGGCGCCCAAUUUUGGGUUUCAGUUGCUGUUUCCAGCUAUGGGGAGAGGUAAGGUU  180
    T W R P I L G F S C C F Q L W G E V R L
                                     M G R G K V
181  AAACCAAAACCGUAAAUCGACGGGUGACAAUUCGAAUGUUGUACUAUGAUUAGAGCUGGA  240
    N Q T V N R R V T I R M L L L
    K P N R K S T G D N S N V V T M I R A G
241  AGCUAUCCUAAGGUCAAUCCGACUCCAACGUGGGUCAGAGCCAUACCUUUCGAAGUGUCA  300
    S Y P K V N P T P T W V R A I P F E V S
301  GUUCAAUCUGGUUUAUUGCUUUUAAAGUACCGGUCGGGUCACUAAUUUUGGCAAAUUUCCGG  360
    V Q S G I A F K V P V G S L F S A N F R
361  ACAGAUUCCUUUACAAGCGUCACAGUGAUGAGUGUCCGUGCUUGGACCCAGUUAACACCG  420
    T D S F T S V T V M S V R A W T Q L T P
421  CCAGUAAAUGAGUACAGUUUUGUGAGGCUGAAGCCAUUGUUAAGACUGGUGACUCUACU  480
    P V N E Y S F V R L K P L F K T G D S T
481  GAGGAGUUCGAAGGGCGUGCAUCAAACAUACACACAGCUUCUGUAGGGUACAGGAUU  540
    E E F E G R A S N I N T R A S V G Y R I
541  CCAACUAAUUUGCGUCAGAAUACUGUGGCAGCCGACAAUGUAUGCGAAGUAAGAAGUAAC  600
    P T N L R Q N T V A A D N V C E V R S N
601  UGUCGACAAGUCGCCUUGGUUAAUUUGGUGUUUUUAAUGAACCUCGACAUAAAGCCUUU  660
    C R Q V A L V I S C C F N
661  UGGAUCGAAGGUUAAACGAUCCGCUCCUGCGUUGAGCUUGAGGCGGCGUAUCUCUUAUGU  720
721  CAACAGAGACACUUUGGUCUAUGGUUGUAUAAUAGAUAGACUCCCGUUUGCAAGAUU  780
781  AGGGUUAACAGAUUCUUGCCGUUAGUCUGGUUAGCGCGUAACCGGCCUUGAUUUUUGGAU  840
841  AGAUCCAUUGUCCAAUGGCUUUGCCAAUGGAACGCCGACGUGGCUGUAUAAUACGUCGU  900
901  GACAAGUACGAAAUUCUUGUAGUGUUUUUCCCUCCACUAAAUCGAAGGGUUUUGUUUUG  960
961  GUCUUCGCCGAACGCAUACGUUAGUGUGACUACCGUUGUUGGAAACAAGUAAAACAGGAAG  1020
1021 GGGGUUCGAAUCCCUCCCUAACCGCGGGUAAGCGGCCCA  1059

```

FIGURE 1

A

2

```

TMV U5 -----
STMV  CCGUUUGCAAGAUUAGGGUUAACAGAUUUUC---CGUUAGUCUGGUUAG
      **      *** ** * * * * *      **      *** **
TMV U1 UAGUAGAAUUGAUCAGAGGAACCGGAUCUUUAAUUCGGAGCUCUUUCGAG

TMV U5 -----CACAAUCUCCGGCUACUUAG
      *      ***      ***
STMV  CGCGUAAACCGGCCUUGAUUUUAGGAAUAGAUCCAUUGUCCAAUGGCUU--
      **      *** * *** * * * * *      ***      ***
TMV U1 AGCU-----CUUCUGGUUUGGUUUGGACCUCUGGUCCUGCAACUUGA

TMV U5 CUAUUGUUGUGAGAUUUUCCUAAAAUAAAGUCGCUGAAGACUUAAAAUUA
      * * * * * * * * * * * * * * * * * * * * *
STMV  -----UGCCAAUGGAACGCCGACGUGGCUGUAUAAUACG
      * * * * * * * * * * * * * * * * * * * * *
TMV U1 -----GGUAGUCAAGAUGCAUAAUAAUAAACGGAUUGUGCCGUAAUCAC

TMV U5 GGG-UGGCUGAUACCAAAUACAGC--AGUGGUUGUUCGUCCACUUAAUA
      * * * * * * * * * * * * * * * * * * * * *
STMV  UCGUUGACAAGUACG-AAAUCUUGUUAGUGUUUUUCCUCCACUUAAUUC
      * * * * * * * * * * * * * * * * * * * * *
TMV U1 ACG-UGGUGCGUACG-AUAACGCA-UAGUGUUUUUCCUCCACUUAAUUC

TMV U5 UAACGAUUGUCAUAUCUGGAUCCUACAGUU-AAAC-CAUGUGAUGGUGUA
      * * * * * * * * * * * * * * * * * * * * *
STMV  GAAGGGUUUUGUUU--UGG-UCUUCGCG----AACGCAUACGUUAGUGUG
      * * * * * * * * * * * * * * * * * * * * *
TMV U1 GAAGGGUUGUGUCU--UGGAUCGCGCGGGUCAAAUGUAUAUGGUUCA-UA

TMV U5 UACUGUGGUUAGG-----CGUAAAACAUCCGAGAGGUUCGAAUCC-U
      * * * * * * * * * * * * * * * * * * * * *
STMV  -ACUACCGU-UGUUCGAAACAAGUAAAACAGGAAGGGGGUUCGAAUCCCU
      * * * * * * * * * * * * * * * * * * * * *
TMV U1 UACAUCGCGCAGG-----CACGUAAUAAAGCGAGGGG-UUCGAAUCCC-

TMV U5 CCCGUAAACCGCCGGUA-GCGGCCCA
      * * * * * * * * * * * * * * * * * * * * *
STMV  CCC-UAACCGCGGGUAAGCGGCCCA
      * * * * * * * * * * * * * * * * * * * * *
TMV U1 CCCGUUACCCCGGUA-GGGGCCCA

```

B

```

BMV  -GUAAAA--UACCAACUAA
      * * * * * * * * * *
STMV  AGUAAAACUUACCAAUCAA
      * * * * * * * * * *
CMV   -GUAA-UCUUACCACUUUC

```

C

```

SPMV GGGUAUUGCA
      ***
STMV  AGUAAAACU
      * * * * *
STNV  AGUAAAGAC

```

FIGURE 2

3

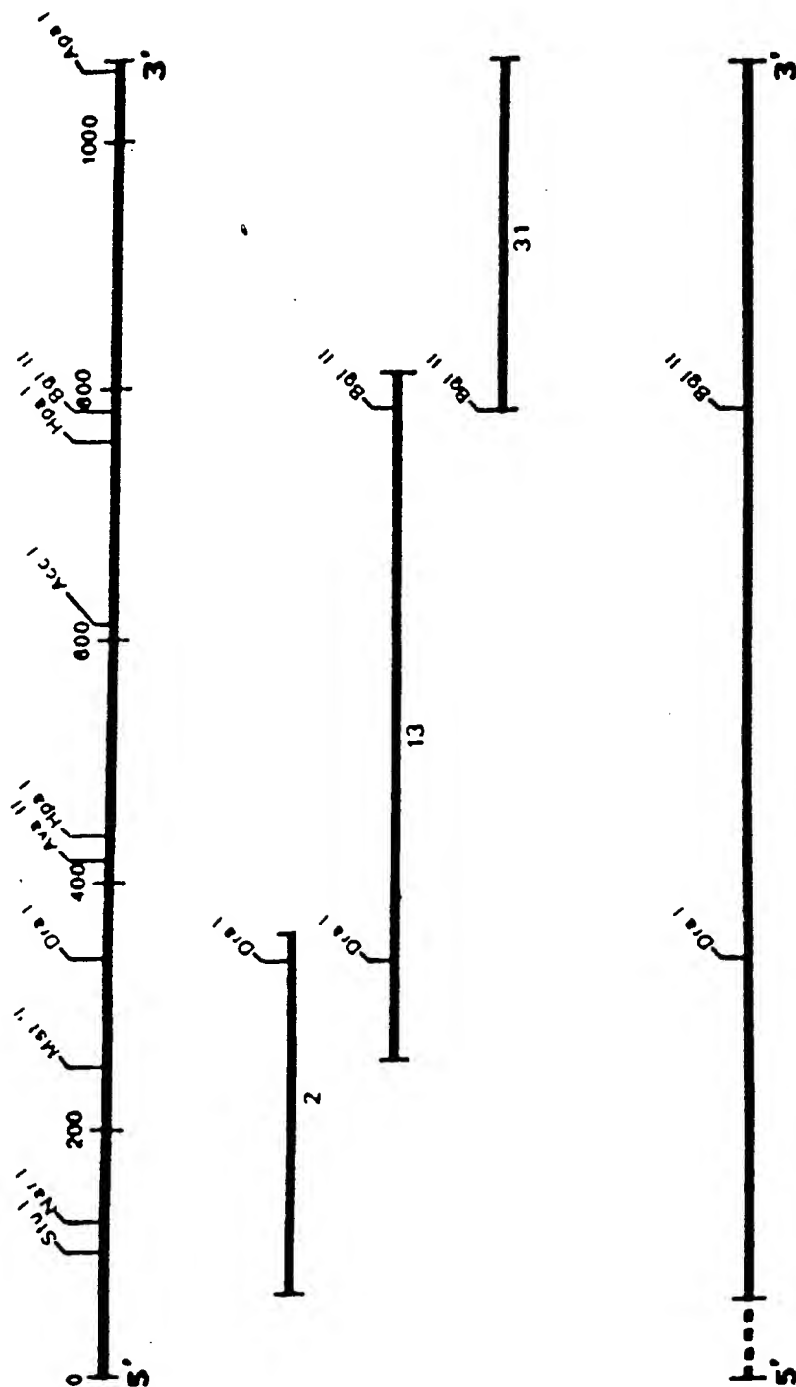
STRATEGY USED TOWARDS CONSTRUCTION OF A  
COMPLETE GENOMIC CLONE FOR STMV

FIGURE 3

# INTERNATIONAL SEARCH REPORT

International Application No PCT/US90/01738

<b>I. CLASSIFICATION F SUBJECT MATTER</b> (if several classification symbols apply, indicate all) <sup>3</sup> According to International Patent Classification (IPC) or to both National Classification and IPC IPC (5): C12P 21/02; C12N 15/00; C07H 21/04 U.S. CL: 435/70.1, 172.3, 320; 536/27																				
<b>II. FIELDS SEARCHED</b> <div style="text-align: center; border-top: 1px solid black; border-bottom: 1px solid black; margin: 5px 0;">Minimum Documentation Searched <sup>4</sup></div> <table style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 30%; text-align: left; border-bottom: 1px solid black;">Classification System</th> <th style="width: 70%; text-align: left; border-bottom: 1px solid black;">Classification Symbols</th> </tr> <tr> <td style="padding: 5px;">U.S.</td> <td style="padding: 5px;">435/69.1, 70.1, 172.3, 317.1, 320 536/27</td> </tr> </table> <div style="text-align: center; border-top: 1px solid black; border-bottom: 1px solid black; margin: 5px 0;">Documentation Searched other than Minimum Documentation to the extent that such documents are included in the fields searched <sup>5</sup></div> <p style="padding: 5px;">USPTO AUTOMATED PATENT SYSTEM: DIALOG FILES: BIOTECH, PATENTS See attachment for search terms.</p>			Classification System	Classification Symbols	U.S.	435/69.1, 70.1, 172.3, 317.1, 320 536/27														
Classification System	Classification Symbols																			
U.S.	435/69.1, 70.1, 172.3, 317.1, 320 536/27																			
<b>III. DOCUMENTS CONSIDERED TO BE RELEVANT</b> <sup>14</sup> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%; text-align: left; padding: 5px;">Category <sup>6</sup></th> <th style="width: 60%; text-align: left; padding: 5px;">Citation of Document, <sup>10</sup> with indication, where appropriate, of the relevant passages <sup>17</sup></th> <th style="width: 30%; text-align: left; padding: 5px;">Relevant to Claim No. <sup>18</sup></th> </tr> </thead> <tbody> <tr> <td style="vertical-align: top; padding: 5px;">X, P Y</td> <td style="vertical-align: top; padding: 5px;">Virology (New York USA), Volume 170, issued May 1989, Mirkov et al, "Nucleotide sequence and translation of satellite tobacco mosaic virus RNA," pages 139-146 (see entire document).</td> <td style="vertical-align: top; padding: 5px;">1-5 1-29</td> </tr> <tr> <td style="vertical-align: top; padding: 5px;">Y, P</td> <td style="vertical-align: top; padding: 5px;">Journal of Cellular Biochemistry (New York, USA) Volume 13 Part D, issued April 1989, Mirkov et al, "Development of satellite tobacco mosaic virus for the expression of heterologous genes in plants," see page 265 (ABSTRACT #M139)</td> <td style="vertical-align: top; padding: 5px;">1-29</td> </tr> <tr> <td style="vertical-align: top; padding: 5px;">Y</td> <td style="vertical-align: top; padding: 5px;">Plant Molecular Biology Reporter (New York, USA), Volume 1, issued 1983, Van Vloten-Doting, "Advantages of multipartite genomes of single stranded RNA plant viruses in nature, for research, and for genetic engineering," pages 55-60 (see entire document).</td> <td style="vertical-align: top; padding: 5px;">1-29</td> </tr> <tr> <td style="vertical-align: top; padding: 5px;">Y</td> <td style="vertical-align: top; padding: 5px;">Journal of general virology (Great Britain), Volume 67, issued 1986, Valverde et al, "Evidence for a satellite RNA associated naturally with the U5 strain and experimentally with the U1 strain of tobacco mosaic virus," pages 1875-1884. (note pages 1879-1881).</td> <td style="vertical-align: top; padding: 5px;">22-29</td> </tr> <tr> <td style="vertical-align: top; padding: 5px;">Y</td> <td style="vertical-align: top; padding: 5px;">Journal of general virology (Great Britain), Volume 68, issued 1987, Valverde et al, "Same properties of isometric virus particles which contain the satellite RNA of tobacco mosaic virus," pages 965-972 (note page 969 and 972).</td> <td style="vertical-align: top; padding: 5px;">1-22</td> </tr> </tbody> </table> <div style="font-size: small; padding: 5px;"> <p><sup>6</sup> Special categories of cited documents: <sup>15</sup></p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>"4" document member of the same patent family</p> </div>			Category <sup>6</sup>	Citation of Document, <sup>10</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>	X, P Y	Virology (New York USA), Volume 170, issued May 1989, Mirkov et al, "Nucleotide sequence and translation of satellite tobacco mosaic virus RNA," pages 139-146 (see entire document).	1-5 1-29	Y, P	Journal of Cellular Biochemistry (New York, USA) Volume 13 Part D, issued April 1989, Mirkov et al, "Development of satellite tobacco mosaic virus for the expression of heterologous genes in plants," see page 265 (ABSTRACT #M139)	1-29	Y	Plant Molecular Biology Reporter (New York, USA), Volume 1, issued 1983, Van Vloten-Doting, "Advantages of multipartite genomes of single stranded RNA plant viruses in nature, for research, and for genetic engineering," pages 55-60 (see entire document).	1-29	Y	Journal of general virology (Great Britain), Volume 67, issued 1986, Valverde et al, "Evidence for a satellite RNA associated naturally with the U5 strain and experimentally with the U1 strain of tobacco mosaic virus," pages 1875-1884. (note pages 1879-1881).	22-29	Y	Journal of general virology (Great Britain), Volume 68, issued 1987, Valverde et al, "Same properties of isometric virus particles which contain the satellite RNA of tobacco mosaic virus," pages 965-972 (note page 969 and 972).	1-22
Category <sup>6</sup>	Citation of Document, <sup>10</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>																		
X, P Y	Virology (New York USA), Volume 170, issued May 1989, Mirkov et al, "Nucleotide sequence and translation of satellite tobacco mosaic virus RNA," pages 139-146 (see entire document).	1-5 1-29																		
Y, P	Journal of Cellular Biochemistry (New York, USA) Volume 13 Part D, issued April 1989, Mirkov et al, "Development of satellite tobacco mosaic virus for the expression of heterologous genes in plants," see page 265 (ABSTRACT #M139)	1-29																		
Y	Plant Molecular Biology Reporter (New York, USA), Volume 1, issued 1983, Van Vloten-Doting, "Advantages of multipartite genomes of single stranded RNA plant viruses in nature, for research, and for genetic engineering," pages 55-60 (see entire document).	1-29																		
Y	Journal of general virology (Great Britain), Volume 67, issued 1986, Valverde et al, "Evidence for a satellite RNA associated naturally with the U5 strain and experimentally with the U1 strain of tobacco mosaic virus," pages 1875-1884. (note pages 1879-1881).	22-29																		
Y	Journal of general virology (Great Britain), Volume 68, issued 1987, Valverde et al, "Same properties of isometric virus particles which contain the satellite RNA of tobacco mosaic virus," pages 965-972 (note page 969 and 972).	1-22																		
<b>IV. CERTIFICATION</b> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; padding: 5px; vertical-align: top;">           Date of the Actual Completion of the International Search <sup>2</sup>   <div style="text-align: center; font-weight: bold; font-size: 1.2em;">15 JUNE 1990</div> </td> <td style="width: 50%; padding: 5px; vertical-align: top;">           Date of Mailing of this International Search Report <sup>3</sup>   <div style="text-align: center; font-weight: bold; font-size: 1.2em;">26 JUL 1990</div> </td> </tr> <tr> <td style="padding: 5px; vertical-align: top;">           International Searching Authority <sup>1</sup>   <div style="text-align: center; font-weight: bold;">ISA/US</div> </td> <td style="padding: 5px; vertical-align: top;">           Signature of Authorized Officer <sup>10</sup>  <div style="text-align: center;">              P. RHODES           </div> </td> </tr> </table> <div style="text-align: right; padding: 5px; font-weight: bold; font-size: 0.8em;">       NGUYEN NGOC HO        INTERNATIONAL DIVISION     </div>			Date of the Actual Completion of the International Search <sup>2</sup>  <div style="text-align: center; font-weight: bold; font-size: 1.2em;">15 JUNE 1990</div>	Date of Mailing of this International Search Report <sup>3</sup>  <div style="text-align: center; font-weight: bold; font-size: 1.2em;">26 JUL 1990</div>	International Searching Authority <sup>1</sup>  <div style="text-align: center; font-weight: bold;">ISA/US</div>	Signature of Authorized Officer <sup>10</sup> <div style="text-align: center;">              P. RHODES           </div>														
Date of the Actual Completion of the International Search <sup>2</sup>  <div style="text-align: center; font-weight: bold; font-size: 1.2em;">15 JUNE 1990</div>	Date of Mailing of this International Search Report <sup>3</sup>  <div style="text-align: center; font-weight: bold; font-size: 1.2em;">26 JUL 1990</div>																			
International Searching Authority <sup>1</sup>  <div style="text-align: center; font-weight: bold;">ISA/US</div>	Signature of Authorized Officer <sup>10</sup> <div style="text-align: center;">              P. RHODES           </div>																			

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category *	Citation of Document, <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No <sup>18</sup>
Y	Virology (New York, USA), Volume 158, issued 1987, Janda et al, "High Efficiency T7 Polymerase Synthesis of Infectious RNA from Cloned Brome Mosaic Virus cDNA and effects of 5' extensions on transcript infectivity," pages 259-262 (entire document)	1-29
Y	Molecular and Cellular Biology (Washington DC, USA), Volume 4, issued 1984, Ahlquist et al, "cDNA cloning and in vitro transcription of the complete Brome mosaic virus genome," pages 2876-2882 (entire document)	1-29
Y	Proceedings National Academy Sciences USA (Washington DC), Volume 81, issued 1984, Ahlquist et al, "multi-component RNA plant virus infection derived from cloned viral cDNA," pages 7066-7070 (entire document and see page 7070, last paragraph).	1-29
Y	EP, A, 0194809 (Ahlquist et al.) 17 September 1986 (entire document but note especially pages 22-23)	1-29
Y	Science (Washington DC, USA), Volume 231, issued 1986, French et al, "Bacterial gene inserted in an engineered RNA virus: efficient expression in monocotyledonous plant cells," pages 1294-1297. (Note last paragraph).	1-29
Y	Virology (New York, USA), Volume 157, issued 1987, Emmelo et al, "expression in Plants of the cloned satellite tobacco necrosis virus genome and of derived insertion mutants," pages 480-487. (entire document).	1-29
Y	EP, A, 0242016 (BAULCOMBE ET AL.) 21 October 1987 (see entire document).	1-29
Y	EP, A, 0278667 (AHLQUIST ET AL.) 09 February 1987 (see entire document).	1-29

PCT/US90/01738

Attachment to Form PCT/ISA/210, Part II.

II. FIELDS SEARCHED TERMS:

STMV, TMV, BMV, vector, RNA virus or viral

Attachment to Form PCT/ISA/210, Part VI.

VI. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

I Claims 1-27, drawn to a recombinant RNA and a DNA vector encoding same and to a method of use to transform plant cells; class 435, subclasses 172.1, 172.3, 317.1 and 320.

II Claims 28-29, drawn to a method of use to make protein; class 435, subclass 69.1.